Generalization of Grover's Algorithm to Multiobject Search in Quantum Computing, Part I: Continuous Time and Discrete Time

Goong Chen[1,2] Stephen A. Fulling[1] Jeesen Chen[3]

Abstract

L. K. Grover's search algorithm in quantum computing gives an optimal, quadratic speedup in the search for a single object in a large unsorted database. In this paper, we generalize Grover's algorithm in a Hilbert-space framework for both continuous and discrete time cases that isolates its geometrical essence to the case where more than one object satisfies the search criterion.

[1] Department of Mathematics, Texas A&M University, College Station, TX 77843-3368. E-mail: gchen@math.tamu.edu and fulling@math.tamu.edu.

[2] Supported in part by Texas A&M University Interdisciplinary Research Grant IRI 99-22.

[3] Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH 45221-0025. E-mail: chenjn@email.uc.edu.

1 Introduction

A quantum computer (QC) is envisaged as a collection of 2-state "quantum bits", or *qubits* (e.g., spin 1/2 particles). Quantum computation does calculations on data densely coded in the entangled states that are the hallmark of quantum mechanics, potentially yielding unprecedented parallelism in computation, as P. Shor's work on factorization [13, 14] proved in 1994. Two years later, L. K. Grover [7] showed that for an unsorted database with N items in storage, it takes an average number of $\mathcal{O}(\sqrt{N})$ searches to locate a single desired object by his quantum search algorithm. If N is a very large number, this is a significant quadratic speedup over the exhaustive search algorithm in a classical computer, which requires an average number of $\frac{N+1}{2}$ searches (see Remark A.1 in Appendix). Even though Grover's algorithm is not exponentially fast (as Shor's is), it has been argued that the wide range of its applicability compensates for this [4]. Furthermore, the quantum speedup of the search algorithm is *indisputable*, whereas for factoring the nonexistence of competitively fast classical algorithms has not yet been proved [1, 2].

Grover's original papers [7, 8] deal with search for a single object. In practical applications, typically more than one item will satisfy the criterion used for searching. In the simplest generalization of Grover's algorithm, the number of "good" items is known in advance (and greater than 1). Here we expound this generalization, along the lines of a treatment of the single-object case by Farhi and Gutmann [6] that makes the Hilbert-space geometry of the situation very clear.

The success of Grover's algorithm and its multiobject generalization is attributable to two main sources:

(i) the notion of amplitude amplication; and

(ii) the dramatic reduction to invariant subspaces of low dimension for the unitary operators involved.

Indeed, the second of these can be said to be responsible for the first: A proper geometrical formulation of the process shows that all the "action" takes place within a *two-dimensional, real* subspace of the Hilbert space of quantum states. Since the state vectors are normalized, the state is confined to a one-dimensional unit circle and (if moved at all) initially has nowhere to go except toward the place where the amplitude for the sought-for state is maximized. This accounts for the robustness of Grover's algorithm — that is, the fact that Grover's original choice of initial state and of the Walsh–Hadamard transformation can be replaced by (almost) any initial state and (almost) any unitary transformation [9, 10, 4].

The notion of amplitude amplification was emphasized in the original works [7, 8, 9] of Grover himself and in those of Boyer, Brassard, Høyer and Tapp [3] and Brassard, Høyer and Tapp [4].

(See also [1, 2].) Dimensional reduction is prominent in the papers by Farhi and Gutmann [6] and Jozsa [10]. We applied dimensional reduction to multiobject search independently of references [3] and [4] and later learned that the same conclusions about multiobject search (and more), in the *discrete time case*, had been obtained there in a different framework.

The rest of the paper is divided into three parts. In §2, we present the continuous-time version of the multiobject search algorithm, and in §3 the discrete-time version. In the Appendix, the computational complexity of the classical random multiobject search algorithm is analyzed and some relevant points on the literature are also made.

2 Continuous Time Quantum Computing Algorithm for Multiobject Search

Farhi and Gutmann [6] first considered quantum computation from a different point of view by regarding it as controlled Hamiltonian (continuous) time evolution of a system. This view is definitely proper, because quantum mechanical systems naturally evolve continuously in time. We inherit their point of view in this section.

Let an unsorted database consist of N objects $\{w_j \mid 1 \le j \le N\}$, and let f be an *oracle* (or Boolean) function such that

$$f(w_j) = \begin{cases} 1, & j = 1, 2, \ldots, \ell, \\ 0, & j = \ell + 1, \ell + 2, \ldots, N. \end{cases} \tag{2.1}$$

Here the ℓ elements $\{w_j \mid 1 \le j \le \ell\}$ are the desired objects of search. However, in general ℓ is *not* explicitly given. Note that the assignment of the "desired" objects to the first ℓ values of the index j is just a convention for the purposes of theoretical discussion; from the point of view of the user of the algorithm, the N objects are in random or unknown order, or perhaps better, have no meaningful ordering whatever. Consider now a Hilbert space $\mathcal{H}$ of dimension N with an orthonormal basis $\mathcal{B} = \{|w_j\rangle \mid 1 \le j \le N\}$; each $|w_j\rangle$ is an eigenstate in the quantum computer representing the object w_j in the database. Denote $L = \text{span}\{|w_j\rangle \mid 1 \le j \le \ell\}$. Here we have adopted the notation of *ket* $|\cdot\rangle$ and *bra* $\langle\cdot|$ in mathematical physics to denote, respectively, vectors and linear functionals in $\mathcal{H}$. Suppose we are given a Hamiltonian H in $\mathcal{H}$ and we are told that H has an eigenvalue $E \ne 0$ on the entire subspace L defined above and all the other eigenvalues are zero. The task is to find an eigenvector $|w_j\rangle$ in L that has eigenvalue E. The task is regarded as completed when a *measurement* of the system shows that it is in the state $|w_j\rangle$ for some $j\colon\ 1 \le j \le \ell$.

Define a linear operator H_L, whose action on a basis element is given by

$$H_L|w_j\rangle = \frac{E}{2}(|w_j\rangle - (-1)^{f(w_j)}|w_j\rangle), \qquad j = 1, 2, \ldots, N. \tag{2.2}$$

Note that here we have only utilized the knowledge of f; no knowledge of the desired search objects $\{|w_j\rangle \mid j = 1, 2, \ldots, \ell\}$ is required or utilized since it is assumed to be hidden in the oracle (black box). Nevertheless, since H_L is a linear operator, through *linear extension* we know that H_L is uniquely defined, and it necessarily has the following unique "explicit" representation

$$H_L = E \sum_{j=1}^{\ell} |w_j\rangle\langle w_j|. \tag{2.3}$$

The explicitness of H_L in (2.3) is somewhat misleading. We need to emphasize that ℓ in (2.3) is *not explicitly* known or given since the only knowledge we have is f in (2.1). For the implementation of the algorithms here as well as in the next section, this does not constitute any problem, however, except that in most applications the determination of and the information about ℓ are important. This becomes a separate class of problems calling *counting* that is studied in [4], which we hope to expound further in a sequel. As in [6], we now add to H_L the "driving Hamiltonian"

$$H_D = E|s\rangle\langle s| \tag{2.4}$$

for some (yet arbitrary) unit vector $|s\rangle \in \mathcal{H}$, $|s\rangle \notin L$. This gives the overall Hamiltonian as

$$H = H_L + H_D. \tag{2.5}$$

Our quantum computer is governed by the Schrödinger equation

$$\begin{cases} i\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle, \quad t > 0, \\ |\psi(0)\rangle = |s\rangle, \end{cases} \tag{2.6}$$

as a continuous-time, controlled Hamiltonian system. Since $(2.6)_1$ is autonomous, the state of the system at time t is given by

$$|\psi(t)\rangle = e^{-iHt}|s\rangle, \qquad t \geq 0, \tag{2.7}$$

where e^{-iHt} is the exponential $N \times N$ (time evolution) matrix.

Define an augmented space $\widetilde{L}$ from L and $|s\rangle$:

$$\widetilde{L} = \operatorname{span}(L \cup \{|s\rangle\}). \tag{2.8}$$

Let $\widetilde{L}^{\perp}$ be the orthogonal complement of $\widetilde{L}$ satisfying $\widetilde{L} \oplus \widetilde{L}^{\perp} = \mathcal{H}$, where $\oplus$ denotes the orthogonal direct sum. With respect to this orthogonal decomposition, we now have our first reduction of dimensionality below.

Proposition 2.1. *Fix $|s\rangle \in \mathcal{H}$, $|s\rangle \notin L$ in (2.6). Let $H, \widetilde{L}$ and $\widetilde{L}^{\perp}$ be defined as above. For any $|w\rangle \in \mathcal{H}$, write $|w\rangle = |v\rangle + |u\rangle \in \widetilde{L} \oplus \widetilde{L}^{\perp}$ according to the orthogonal direct sum. Then $H|w\rangle = H|v\rangle$ and $H|u\rangle = 0$. Consequently, the Hamiltonian H has an associated blockwise decomposition*

$$H = \begin{bmatrix} H_{\widetilde{L}} & \vdots & 0_{12} \\ \cdots & \cdots & \cdots \\ 0_{21} & \vdots & 0_{22} \end{bmatrix} \tag{2.9}$$

where $H_{\widetilde{L}}$ is an invertible $(\ell+1)\times(\ell+1)$ matrix defined on $\widetilde{L}$ such that $H_{\widetilde{L}}|v\rangle = H|v\rangle$ for all $|v\rangle \in \widetilde{L}$, and $0_{12}, 0_{21}$ and 0_{22} are, respectively, $(\ell+1)\times(N-\ell-1)$, $(N-\ell-1)\times(\ell+1)$ and $(N-\ell-1)\times(N-\ell-1)$ zero matrices.

Proof. Straightforward verification. □

Corollary 2.2. *Fix $|s\rangle \in \mathcal{H}, |s\rangle \notin L$ in (2.6). Let $H, \widetilde{L}$ an $\widetilde{L}^{\perp}$ be defined as above. Then the state of the solution of (2.6) at time $t, |\psi(t)\rangle$, has zero component in $\widetilde{L}^{\perp}$ for all $t > 0$.*

Proof. The action of the evolution dynamics e^{-iHt} on the invariant subspace $\widetilde{L}^{\perp}$ is, by (2.9), $e^{-i0_{22}t} = e^0 = \boldsymbol{I}_{\widetilde{L}^{\perp}}$, the identity operator on $\widetilde{L}^{\perp}$. Since the component of $|s\rangle$ in $\widetilde{L}^{\perp}$ is the zero vector, the action of $\boldsymbol{I}_{N-\ell-1}$ (the $(N-\ell-1)\times(N-\ell-1)$ identity matrix) on it remains zero for all $t > 0$. □

By the properties obtained above, we need to fix our attention only on $H_{\widetilde{L}}$ defined on $\widetilde{L}$. By abuse of notation, we still write H instead of $H_{\widetilde{L}}$ on $\widetilde{L}$.

Proposition 2.3 (Matrix representation of $\boldsymbol{H}$ on $\widetilde{\boldsymbol{L}}$). *Under the same assumptions as in Prop. 2.1, define $x_i = \langle s|w_i\rangle$ for $i = 1, 2, \ldots, \ell$. Let*

$$|r\rangle = \frac{1}{C_r}\left(|s\rangle - \sum_{i=1}^{\ell} x_i|w_i\rangle\right), \qquad C_r \equiv \sqrt{1 - \sum_{i=1}^{\ell}|x_i|^2}. \tag{2.10}$$

Then $\{|w_1\rangle, \ldots, |w_\ell\rangle, |r\rangle\}$ forms an orthonormal basis of $\widetilde{L}$ with respect to which H admits the matrix representation

$$H = E[H_{ij}]_{(\ell+1)\times(\ell+1)}; \quad H_{ij} = \begin{cases} x_j\bar{x}_i, & 1 \le i,j \le \ell, i \ne j, \\ 1 + |x_j|^2, & 1 \le i,j \le \ell, i = j, \\ (\delta_{j,\ell+1}x_j + \delta_{i,\ell+1}\bar{x}_i)C_r, & i = \ell+1 \text{ or } j = \ell+1, i \ne j, \\ C_r^2, & i = j = \ell+1. \end{cases} \tag{2.11}$$

Proof. Solve (2.10) for $|s\rangle$:

$$|s\rangle = \sum_{i=1}^{\ell} x_i|w_i\rangle + C_r|r\rangle. \tag{2.12}$$

Substituting (2.12) into (2.3) and (2.4), we obtain (2.11) in bra-ket form. □

The exponential matrix function e^{-iHt} on $\widetilde{L}$ based upon the representation can be obtained, e.g., by

$$e^{-iHt} = \sum_{k=0}^{\infty} \frac{t^k}{k!}(-iH)^k \tag{2.13}$$

or

$$e^{-iHt} = \frac{1}{2\pi i} \oint_{\mathcal{C}} (\zeta \boldsymbol{I}_{\ell+1} + iH)^{-1} e^{\zeta t} d\zeta \tag{2.14}$$

where $\mathcal{C}$ is a simple closed curve in $\mathbb{C}$ enclosing all the values ζ such that the $(\ell+1) \times (\ell+1)$ matrix $\zeta \boldsymbol{I}_{\ell+1} + iH$ is not invertible. However, since ℓ can be arbitrarily large, it is a highly nontrivial task (if possible at all) to calculate e^{-iHt} explicitly based on (2.13), (2.14) or any other known techniques.

It turns out that the above difficulty can be bypassed if $|s\rangle$ is chosen in an ingenious way that can effect another reduction of dimensionality. (Nevertheless, we again call attention to the fact that any choice of $|s\rangle$ must be independent of any knowledge of $|w_i\rangle$, for $i = 1, 2, \ldots, \ell$.) The choice by Grover [7, 8] is

$$|s\rangle = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} |w_j\rangle, \tag{2.15}$$

implementable and obtainable on the QC through an application of the Walsh-Hadamard transformation on all the qubits; $|s\rangle$ is a superposition state with the same amplitude in all eigenstates. With this choice of $|s\rangle$, we now have

$$x_i = x \equiv 1/\sqrt{N}, i = 1, 2, \ldots, \ell; \quad C_r^2 = 1 - (\ell/N), \quad |r\rangle = \frac{1}{\sqrt{1-(\ell/N)}} \left(|s\rangle - \frac{1}{\sqrt{N}} \sum_{i=1}^{\ell} |w_i\rangle \right) \tag{2.16}$$

in (2.11) and (2.10).

Theorem 2.4 (A two-dimensional invariant subspace for the Hamiltonian $\boldsymbol{H}$). *Let $|s\rangle$ be given as in (2.15). Denote*

$$\mathcal{V} = \{ |v\rangle \in \widetilde{L} \;\Big|\; |v\rangle = a \sum_{i=1}^{\ell} |w_i\rangle + b|r\rangle; \quad a, b \in \mathbb{C} \}. \tag{2.17}$$

Then $\mathcal{V}$ is a invariant two-dimensional subspace of H in $\widetilde{L}$ such that

(1) $r, s, \in \mathcal{V}$;

(2) $H(\mathcal{V}) = \mathcal{V}$.

Proof. It is obvious that (1) holds. To see (2), we have

$$\begin{aligned}
H|v\rangle &\equiv H\left[a\sum_{i=1}^{\ell}|w_i\rangle + b|r\rangle\right]\\
&= a\sum_{i=1}^{\ell}H|w_i\rangle + bH|r\rangle\\
&= E\Bigg\{a\sum_{i=1}^{\ell}\Bigg[(1+x^2)|w_i\rangle + x^2\sum_{\substack{j=1\\ j\neq i}}^{\ell}|w_j\rangle + x\sqrt{1-\ell x^2}|r\rangle\Bigg]\\
&\quad + b\left[x\sqrt{1-\ell x^2}\sum_{j=1}^{\ell}|w_j\rangle + (1-\ell x^2)|r\rangle\right]\Bigg\}\\
&= E\Bigg\{a\sum_{i=1}^{\ell}\left[|w_i\rangle + x^2\sum_{j=1}^{\ell}|w_j\rangle\right]\\
&\quad + bx\sqrt{1-\ell x^2}\sum_{j=1}^{\ell}|w_j\rangle + [a\ell x\sqrt{1-\ell x^2} + b(1-\ell x^2)]|r\rangle\Bigg\}\\
&= E\left\{(a + a\ell x^2 + bx\sqrt{1-\ell x^2}\sum_{i=1}^{\ell}|w_i\rangle + [a\ell x\sqrt{1-\ell x^2} + b(1-\ell x^2)]|r\rangle\right\} \in \mathcal{V}. \qquad (2.18)
\end{aligned}$$

□

Corollary 2.5. *Define*

$$|\widetilde{w}\rangle = \frac{1}{\sqrt{\ell}}\sum_{i=1}^{\ell}|w_i\rangle. \qquad (2.19)$$

Then $\{|\widetilde{w}\rangle, |r\rangle\}$ *forms an orthonormal basis for* $\mathcal{V}$ *such that with respect to this basis,* H *admits the matrix representation*

$$H = E\begin{bmatrix} 1+\frac{\ell}{N} & \frac{\sqrt{\ell(N-\ell)}}{N} \\ \frac{\sqrt{\ell(N-\ell)}}{N} & 1-\frac{\ell}{N}\end{bmatrix} \qquad (2.20)$$

with

$$e^{-iHt} = e^{-iEt}\begin{bmatrix}\cos(Eyt) - iy\sin(Eyt) & -\sqrt{1-y^2}i\sin(Eyt)\\ -\sqrt{1-y^2}i\sin(Eyt) & \cos(Eyt) + iy\sin(Eyt)\end{bmatrix}, y \equiv \sqrt{\ell/N}. \qquad (2.21)$$

Proof. The representation (2.20) follows easily from (2.18).

To calculate e^{-iHt}, write

$$H = E\begin{bmatrix}1+y^2 & y\sqrt{1-y^2}\\ y\sqrt{1-y^2} & 1-y^2\end{bmatrix}, \qquad y \equiv \sqrt{\ell/N},$$

analogous to [6, (8)], and apply (2.14) to obtain (2.21); or apply (2.13) and the properties of the $SU(2)$ generators commonly used in quantum mechanics [12, (XIII.84), p. 546]. □

Since $\mathcal{V}$ is invariant under H and H is a self-adjoint matrix, the orthogonal complement $\mathcal{V}^\perp$ of $\mathcal{V}$ is also an invariant subspace of H. The precise action of H or e^{iHt} on $\mathcal{V}^\perp$ does not seem to be describable in simple terms. However, this knowledge is not needed as we now have the explicit form of the solution available below.

Corollary 2.6. *The solution $\psi(t)$ of (2.6) given (2.15) is*

$$\psi(t) = e^{-iEt}\{[y\cos(Eyt) - i\sin(Eyt)]|\widetilde{w}\rangle + \sqrt{1-y^2}\cos(Eyt)|r\rangle\}, \quad t > 0, \tag{2.22}$$

where $y = \sqrt{\ell/N}$.

Proof. Use (2.21) and (2.16). □

Note that (2.22) has the same form as [6, (10)]. Since $|\widetilde{w}\rangle$ and $|r\rangle$ are unit, mutually orthogonal vectors, the probability of reaching the state $|\widetilde{w}\rangle$ at time t is

$$P(t) = \sin^2(Eyt) + y^2\cos^2(Eyt), \qquad y = \sqrt{\ell/N}. \tag{2.23}$$

At time $T \equiv \pi/(2Ey)$, the probability is 1. By (2.19), if we make a measurement of $|\psi(t)\rangle$ at $t = T$, we will obtain any one of the eigenstates $|w_i\rangle$, $i = 1, 2, \ldots, \ell$, with equal probability $1/\ell$. Therefore the task of search is completed (*with probability 1*), requiring a total time duration

$$T = \pi/(2Ey) = \frac{\pi}{2E}\sqrt{\frac{N}{\ell}}. \tag{2.24}$$

Formula (2.23) manifests the notion of *amplitude amplication* because the amplitude of $|\psi(t)\rangle$ along $|r\rangle$, $\sqrt{1-y^2}\cos(Eyt)$, is steadily decreasing in magnitude as a function of $t \in [0, T]$, implying that $P(t)$ in (2.23) is increasing for $t \in [0, T]$.

Next, let us address the optimality of the search algorithm as given above. We assume that N/ℓ is a large number. We show that the above generalized Grover-Farhi-Gutmann algorithm for multiobject search is time-optimal within the order $\mathcal{O}(\sqrt{N/\ell})$. In contrast with the classical random search which requires in average $(N+1)/(\ell+1)$ searches (see (A.6) in the Appendix), again we see that there is a quadratic speedup.

The idea of proof follows by combining those in [3] and [6]. Let the Hamiltonian H_L be given as in (2.3). We wish to add a somewhat arbitrary (generally, time-dependent) driving Hamiltonian $H_D(t)$ to H_L so that the terminal state (at time $\widetilde{T}$) is in L (which, after a measurement, becomes one of the eigenstates $|w_1\rangle, |w_2\rangle, \ldots, |w_\ell\rangle$). Our objective is to find a lower bound on $\widetilde{T}$. Of course, $H_D(t)$ and $\widetilde{T}$ must be independent of L, since they are part of the algorithm prescribed for determining L.

Let $|w_I\rangle \in \mathcal{H}$ be an arbitrary initial state such that $\langle w_I|w_I\rangle = 1$. Let $\psi_L(t)$ denote the solution of the Schrödinger equation

$$\begin{cases} i\dfrac{d}{dt}|\psi_L(t)\rangle = (H_L + H_D(t))|\psi_L(t)\rangle, \qquad 0 < t \le \widetilde{T}, \\ |\psi_L(0)\rangle = |w_I\rangle, \qquad \text{(initial condition)} \\ |\psi_L(\widetilde{T})\rangle = \displaystyle\sum_{i=1}^{\ell} \alpha_i |w_i\rangle \in L, \qquad \text{(terminal condition).} \end{cases} \tag{2.25}$$

Note that $\alpha_i \in \mathbb{C}$ for $i = 1, 2, \dots, \ell$ and

$$\sum_{i=1}^{\ell} |\alpha_i|^2 = 1 \tag{2.26}$$

because the evolution process is unitary at any $t \in (0, \widetilde{T}]$. On the other hand, let $|\psi(t)\rangle$ evolve with $H_D(t)$:

$$\begin{cases} i\dfrac{d}{dt}|\psi(t)\rangle = H_D(t)|\psi(t)\rangle, \quad 0 < t \le \widetilde{T}, \\ |\psi(0)\rangle = |w_I\rangle. \end{cases} \tag{2.27}$$

Lemma 2.7. *Assume that $\widetilde{N} \equiv N/\ell$ is an integer. Let the orthonormal basis $\mathcal{B} = \{|w_i\rangle \mid 1 \le i \le N\}$ be grouped into $\widetilde{N}$ disjoint subsets*

$$\mathcal{B} = \dot{\bigcup}_{k=1}^{\widetilde{N}} B_k \tag{2.28}$$

where each $B_k \equiv \{|w_{i,k}\rangle \mid i = 1, 2, \dots, \ell\}$ contains exactly ℓ orthonormal basis elements. Then we have

$$\begin{aligned} 2\widetilde{N} - 2\sqrt{\widetilde{N}} &\le \sum_{k=1}^{\widetilde{N}} \langle \psi_{L_k}(\widetilde{T}) - \psi(\widetilde{T}) | \psi_{L_k}(\widetilde{T}) - \psi(\widetilde{T})\rangle \\ &\le 2E\sqrt{\widetilde{N}}\ \widetilde{T}. \end{aligned} \tag{2.29}$$

where $|\psi_{L_k}(t)\rangle$ is the solution $|\psi_L(t)\rangle$ of (2.25) with $L = L_k = span\ B_k$ in both the differential equation and the terminal condition. Consequently,

$$\widetilde{T} \ge (1 - \varepsilon_{\widetilde{N}})\frac{\widetilde{N}^{1/2}}{E} = \frac{1 - \varepsilon_{\widetilde{N}}}{E}\sqrt{\frac{N}{\ell}}, \tag{2.30}$$

where $\varepsilon_{\widetilde{N}} = \widetilde{N}^{-1/2} \to 0$ as $\widetilde{N} \to \infty$.

Proof. As in [6, (1.8)], we have

$$\langle \psi_{L_k}(\widetilde{T}) - \psi(\widetilde{T}) | \psi_{L_k}(\widetilde{T}) - \psi(\widetilde{T})\rangle = 2 - \sum_{i=1}^{\ell} [\langle \alpha_{i,k} w_{i,k} | \psi(\widetilde{T})\rangle + \langle \psi(\widetilde{T}) | \alpha_{i,k} w_{i,k}\rangle], \tag{2.31}$$

and, therefore

$$\sum_{k=1}^{\widetilde{N}} \langle \psi_{L_k}(\widetilde{T}) - \psi(\widetilde{T}) | \psi_{L_k}(\widetilde{T}) - \psi(\widetilde{T})\rangle = 2\widetilde{N} - \sum_{k=1}^{\widetilde{N}} \left[\Big\langle \sum_{i=1}^{\ell} \alpha_{i,k} w_{i,k} \Big| \psi(\widetilde{T}) \Big\rangle + \Big\langle \psi(\widetilde{T}) \Big| \sum_{i=1}^{\ell} \alpha_{i,k} w_{i,k} \Big\rangle \right].$$

Let $\mathbb{P}_{L_k}\colon\ \mathcal{H} \to L_k$ be the orthogonal projection of $\mathcal{H}$ onto L_k. Then

$$\sum_{k=1}^{\widetilde{N}} \left|\left\langle \sum_{i=1}^{\ell} \alpha_{i,k} w_{i,k} \middle| \psi(\widetilde{T})\right\rangle\right|^2 \leq \sum_{k=1}^{\widetilde{N}} |\mathbb{P}_{L_k}|\psi(\widetilde{T})\rangle\|^2 \leq 1, \tag{2.32}$$

from which, we apply the Cauchy-Schwarz inequality and obtain

$$\sum_{k=1}^{\widetilde{N}} \left|\left\langle \sum_{i=1}^{\ell} \alpha_{i,k} w_{i,k} \middle| \psi(\widetilde{T})\right\rangle + \left\langle \psi(\widetilde{T}) \middle| \sum_{i=1}^{\ell} \alpha_{i,k} w_{i,k}\right\rangle\right| \leq 2\widetilde{N}^{1/2}. \tag{2.33}$$

Combining (2.31) and (2.33), we have established the left half of the inequality (2.29).

Next, mimicking [6, (19)–(21)], we have

$$\begin{aligned}
\frac{d}{dt}[\langle \psi_{L_k}(t) - \psi(t)|\psi_{L_k}(t) - \psi(t)\rangle] &= 2\ \mathrm{Im}\langle \psi_{L_k}(t)|H_{L_k}|\psi(t)\rangle \\
&\leq 2|\langle \psi_{L_k}(t)|H_{L_k}|\psi(t)\rangle| \\
&\leq 2\|H_{L_k}|\psi(t)\rangle\| = 2E\left[\sum_{i=1}^{\ell} |\langle w_{i,k}|\psi(t)\rangle|^2\right]^{1/2} \\
&= 2E|\mathbb{P}_{L_k}|\psi(t)\rangle|, \\
\frac{d}{dt}\sum_{k=1}^{\widetilde{N}}[\langle \psi_{L_k}(t) - \psi(t)|\psi_{L_k}(t) - \psi(t)\rangle] &\leq 2E\sum_{k=1}^{\widetilde{N}} |\mathbb{P}_{L_k}|\psi(t)\rangle|,
\end{aligned}$$

and from $\sum\limits_{k=1}^{\widetilde{N}} |\mathbb{P}_{L_k}|\psi(t)\rangle|^2 = 1$, by an application of the Cauchy-Schwarz inequality again, we obtain

$$\frac{d}{dt}\sum_{k=1}^{\widetilde{N}}[\langle \psi_{L_k}(t) - \psi(t)|\psi_{L_k}(t) - \psi(t)\rangle] \leq 2E\widetilde{N}^{1/2}. \tag{2.34}$$

Integrating (2.34) from 0 to $\widetilde{T}$, noting that $|\psi_{L_k}(0) - \psi(0)\rangle = 0$, we have verified the right half of inequality (2.29). □

In general, N/ℓ is not necessarily an integer. Therefore, the disjoint union (2.28) is not always possible. Define $\widetilde{N} = [N/\ell]$, namely, the integral part of the rational number N/ℓ. Then

$$\frac{N}{\ell} = \widetilde{N} + \delta, \quad \text{where} \quad 0 \leq \delta < 1.$$

Then we can rewrite (2.28) as

$$\mathcal{B} = \dot{\bigcup}_{k=1}^{\widetilde{N}} B_k \cup R, \tag{2.35}$$

where $\dot{\bigcup}_{k=1}^{\widetilde{N}} B_k$ is an arbitrary collection of disjoint sets of ℓ orthonormal basis elements containing a total of $\widetilde{N}\ell$ of them, and R is the remaining set of orthonormal basis elements with cardinality $\ell\delta(< \ell)$. The proof of Lemma 2.7 extends to this case except for some tedious details of bookkeeping concerning the short set R, which we omit. We therefore have arrived at the following order of time-optimality for the continuous-time generalized Grover algorithm for multiobject search.

Theorem 2.8. *Assume that N/ℓ is large. Then it requires at least*

$$\widetilde{T} = \frac{1-\varepsilon_{N/\ell}}{E}\sqrt{\frac{N}{\ell}}, \quad \varepsilon_{N/\ell} > 0, \quad \varepsilon_{N/\ell} = \mathcal{O}((N/\ell)^{-1/2}),$$

time duration in average for the driven general Hamiltonian system (2.27) to reach the subspace L. □

3 Discrete Time Case: Straightforward Generalization of Grover's Algorithm to Multiobject Search

In this section we generalize Grover's search algorithm in its original form [7, 8] to the situation where the number of objects satisfying the search criterion is greater than 1. We are considering the discrete time case here, which may be regarded as a discrete-time sampled system of the continuous-time case treated in the preceding section. Unlike the continuous-time case, there have been a relatively rich literature studying the generalization of Grover's discrete-time algorithm to multiobject search (see [1, 2, 3, 4] and the references therein). Our presentation below gives more formalized arguments than those earlier contributions, provides a clearer Hilbert space framework, and settles a relevant question contained in [3].

Let the database $\{w_i \mid i = 1, 2, \ldots, N\}$, orthonormal eigenstates $\{|w_i\rangle \mid i = 1, 2, \ldots, N\}$ and the oracle function f be the same as given at the beginning of §2. The definitions of $\mathcal{H}, L$ and ℓ remain the same.

Define a linear operation in terms of the oracle function f as follows:

$$I_L|w_j\rangle = (-1)^{f(w_j)}|w_j\rangle, \qquad j = 1, 2, \ldots, N. \tag{3.1}$$

Then since I_L is linear, the extension of I_L to the entire space $\mathcal{H}$ is unique, with an "explicit" representation

$$I_L = \boldsymbol{I} - 2\sum_{j=1}^{\ell} |w_j\rangle\langle w_j|, \tag{3.2}$$

where $\boldsymbol{I}$ is the identity operator on $\mathcal{H}$. I_L is the operator of *rotation (by π) of the phase* of the subspace L. Note again that the explicitness of (3.2) is misleading because explicit knowledge of $\{|w_j\rangle \mid 1 \le j \le \ell\}$ and ℓ in (3.2) is not available. Nevertheless, (3.2) is a well-defined (*unitary*) operator on $\mathcal{H}$ because of (3.1). (Unitarity is a requirement for all operations in a QC.)

We now define $|s\rangle$ as in (2.15). Then

$$|s\rangle = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}|w_i\rangle = \frac{1}{\sqrt{N}}\sum_{i=1}^{\ell}|w_i\rangle + \sqrt{\frac{N-\ell}{N}}|r\rangle; \text{ see (2.16) for } |r\rangle. \tag{3.3}$$

Now, define another operator, *the inversion about average operation*, just as in Grover [7, 8]:

$$I_s = \boldsymbol{I} - 2|s\rangle\langle s|. \tag{3.4}$$

Note that I_s in (3.4) is unitary and hence quantum mechanically admissible. I_s is *explicitly known*, constructible with the so-called Walsh-Hadamard transformation.

Lemma 3.1. *Let $\widetilde{L}$ be defined as in (2.8). Then $\{|w_i\rangle, |r\rangle \mid i = 1, 2, \ldots, \ell\}$ forms an orthonormal basis of $\widetilde{L}$. The orthogonal direct sum $\mathcal{H} = \widetilde{L} \oplus \widetilde{L}^\perp$ is an orthogonal invariant decomposition for both operators $I_{\widetilde{L}}$ and I_s. Furthermore,*

(i) The restriction of I_s on $\widetilde{L}$ admits a unitary matrix representation with respect to the orthonormal basis $\{|w_1\rangle, |w_2\rangle, \ldots, |w_\ell\rangle, |r\rangle\}$:

$$\begin{aligned} A &= [a_{ij}]_{(\ell+1)\times(\ell+1)}, \\ a_{ij} &= \begin{cases} \delta_{ij} - \dfrac{2}{N}, & 1 \le i, j \le \ell, \\ -\dfrac{2\sqrt{N-\ell}}{N}(\delta_{i,\ell+1} + \delta_{j,\ell+1}), & i = \ell+1 \text{ or } j = \ell+1, i \ne j, \\ \dfrac{2\ell}{N} - 1, & i = j = \ell + 1. \end{cases} \end{aligned} \tag{3.5}$$

(ii) The restriction of I_s of $\widetilde{L}^\perp$ is $\mathbb{P}_{\widetilde{L}^\perp}$, the orthogonal projection operator on $\widetilde{L}^\perp$. Consequently, $I_s|_{\widetilde{L}^\perp} = \boldsymbol{I}_{\widetilde{L}^\perp}$, where $\boldsymbol{I}_{\widetilde{L}^\perp}$ is the identity operator on $\widetilde{L}^\perp$.

Proof. We have, from (3.3) and (3.4),

$$\begin{aligned} I_s &= \boldsymbol{I} - 2\left[\frac{1}{\sqrt{N}}\sum_{i=1}^{\ell}|w_i\rangle + \sqrt{\frac{N-\ell}{N}}|r\rangle\right]\left[\frac{1}{\sqrt{N}}\sum_{j=1}^{\ell}\langle w_j| + \sqrt{\frac{N-\ell}{N}}\langle r|\right] \\ &= \left[\sum_{i=1}^{\ell}|w_i\rangle\langle w_i\rangle + |r\rangle\langle r| + \mathbb{P}_{\widetilde{L}^\perp}\right] - \left\{\frac{2}{N}\sum_{i=1}^{\ell}\sum_{j=1}^{\ell}|w_i\rangle\langle w_j|\right. \\ &\quad \left. + \frac{2\sqrt{N-\ell}}{N}\left[\sum_{i=1}^{\ell}(|w_i\rangle\langle r| + |r\rangle\langle w_i|)\right] + 2\left(\frac{N-\ell}{N}\right)|r\rangle\langle r|\right\} \\ &= \sum_{i=1}^{\ell}\sum_{j=1}^{\ell}\left(\delta_{ij} - \frac{2}{N}\right)|w_i\rangle\langle w_j| - \frac{2\sqrt{N-\ell}}{N}\left[\sum_{i=1}^{\ell}(|w_i\rangle\langle r| + |r\rangle\langle w_i|)\right] \\ &\quad + \left(\frac{2\ell}{N} - 1\right)|r\rangle\langle r| + \mathbb{P}_{\widetilde{L}^\perp}. \end{aligned} \tag{3.6}$$

The conclusion follows. □

The generalized "Grover's search engine" for multiobject search is now defined as

$$U = -I_s I_L. \tag{3.7}$$

Lemma 3.2. *The orthogonal direct sum $\mathcal{H} = \widetilde{L} \oplus \widetilde{L}^\perp$ is an invariant decomposition for the unitary operator U, such that the following holds:*

(1) With respect to the orthonormal basis $\{|w_1\rangle, \dots, |w_\ell\rangle, |r\rangle\}$ *of* $\widetilde{L}, U$ *admits a unitary matrix representation*

$$U|_{\widetilde{L}} = [u_{ij}]_{(\ell+1)\times(\ell+1)},$$
$$u_{ij} = \begin{cases} \delta_{ij} - \dfrac{2}{N}, & 1 \le i, j \le \ell, \\ \dfrac{2\sqrt{N-\ell}}{N}(\delta_{j,\ell+1} - \delta_{i,\ell+1}), & i = \ell+1 \text{ or } j = \ell+1, i \neq j, \\ 1 - \dfrac{2\ell}{N}, & i = j = N+1. \end{cases} \tag{3.8}$$

(2) The restriction of U *on* $\widetilde{L}^\perp$ *is* $-\mathbb{P}_{\widetilde{L}^\perp} = -\boldsymbol{I}_{\widetilde{L}^\perp}$.

Proof. Substituting (3.2) and (3.6) into (3.7) and simplifying, we obtain

$$\begin{aligned} U &= -I_s I_L = \cdots \text{(simplification)} \\ &= \sum_{i=1}^{\ell}\sum_{j=1}^{\ell}\left(\delta_{ij} - \frac{2}{N}\right)|w_i\rangle\langle w_j| + \frac{2\sqrt{N-\ell}}{N}\sum_{i=1}^{\ell}(|w_i\rangle\langle r| - |r\rangle\langle w_i|) \\ &\quad + \left(1 - \frac{2\ell}{N}\right)|r\rangle\langle r| - \mathbb{P}_{\widetilde{L}^\perp}. \end{aligned}$$

The proof follows. □

Lemmas 3.1 and 3.2 above effect a reduction of the problem to an invariant subspace $\widetilde{L}$, just as Prop. 2.1 did. However, $\widetilde{L}$ is an $(\ell+1)$-dimensional subspace where ℓ may also be large. Another reduction of dimensionality is needed to further simplify the operator U.

Proposition 3.3. *Define* $\mathcal{V}$ *as in (2.17). Then* $\mathcal{V}$ *is an invariant two-dimensional subspace of* U *such that*

(1) $r, s \in \mathcal{V}$*;*

(2) $U(\mathcal{V}) = \mathcal{V}$.

Proof. Straightforward verification. □

Let $|\widetilde{w}\rangle$ be defined as in (2.19). Then as in §2, $\{|\widetilde{w}\rangle, |r\rangle\}$ forms an orthonormal basis of $\mathcal{V}$. We have the second reduction, to dimensionality 2.

Theorem 3.4. *With respect to the orthonormal basis* $\{|\widetilde{w}\rangle, |r\rangle\}$ *in the invariant subspace* $\mathcal{V}, U$ *admits the unitary matrix representation*

$$U = \begin{bmatrix} \frac{N-2\ell}{N} & \frac{2\sqrt{\ell(N-\ell)}}{N} \\ -\frac{2\sqrt{\ell(N-\ell)}}{N} & \frac{N-2\ell}{N} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}, \theta \equiv \sin^{-1}\left(\frac{2\sqrt{\ell(N-\ell)}}{N}\right). \tag{3.9}$$

Proof. Use the matrix representation (3.8) and (2.19). □

Since $|s\rangle \in \mathcal{V}$, we can calculate $U^m|s\rangle$ efficiently using (3.9):

$$\begin{aligned}
U^m|s\rangle &= U^m\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{\ell}|w_i\rangle + \sqrt{\frac{N-\ell}{N}}|r\rangle\right) \qquad \text{(by (3.3))}\\
&= U^m\left(\frac{\ell}{\sqrt{N}}|\widetilde{w}\rangle + \sqrt{\frac{N-\ell}{N}}|r\rangle\right)\\
&= \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}^m \begin{bmatrix}\frac{\ell}{\sqrt{N}}\\ \sqrt{\frac{N-\ell}{N}}\end{bmatrix}\\
&= \begin{bmatrix}\cos(m\theta+\alpha)\\ \sin(m\theta+\alpha)\end{bmatrix}, \quad \left(\alpha \equiv \cos^{-1}\frac{\ell}{\sqrt{N}}\right) \qquad (3.10)\\
&= \cos(m\theta+\alpha)\cdot|\widetilde{w}\rangle + \sin(m\theta+\alpha)\cdot|r\rangle.
\end{aligned}$$

Thus, the probability of reaching the state $|\widetilde{w}\rangle$ after m iterations is

$$P_m = \cos^2(m\theta+\alpha); \qquad (3.11)$$

cf. (2.23) in the continuous-time case. If $\ell \ll N$, then α is close to $\pi/2$ and, therefore, (3.11) is an increasing function of m initially. This again manifests the notion of amplitude amplification. This probability P_m is maximized if $m\theta + \alpha = \pi$, implying

$$m = \left[\frac{\pi-\alpha}{\theta}\right] = \text{the integral part of } \frac{\pi-\alpha}{\theta}.$$

When ℓ/N is small, we have

$$\begin{aligned}
\theta &= \sin^{-1}\left(\frac{2\sqrt{\ell(N-\ell)}}{N}\right)\\
&= \sin^{-1}\left(2\sqrt{\frac{\ell}{N}}\left[1-\frac{1}{2}\frac{\ell}{N}-\frac{1}{8}\left(\frac{\ell}{N}\right)^2 \pm\cdots\right]\right)\\
&= 2\sqrt{\frac{\ell}{N}} + \mathcal{O}((\ell/N)^{3/2}),\\
\alpha &= \cos^{-1}\frac{\ell}{\sqrt{N}} = \frac{\pi}{2} - \left[\frac{\ell}{\sqrt{N}} + \mathcal{O}((\ell/N^{1/2})^3)\right].
\end{aligned}$$

Therefore

$$\begin{aligned}
m &\approx \frac{\pi - \left\{\frac{\pi}{2} - \left[\frac{\ell}{\sqrt{N}} + \mathcal{O}((\ell/N^{1/2})^3)\right]\right\}}{2\sqrt{\frac{\ell}{N}} + \mathcal{O}((\ell/N)^{3/2})}\\
&= \frac{\pi}{4}\sqrt{\frac{N}{\ell}}\left[1 + \mathcal{O}\left(\frac{\ell}{N}\right)\right]. \qquad (3.12)
\end{aligned}$$

Corollary 3.5. *The generalized Grover's algorithm for multiobject search with operator U given by (3.7) has success probability $P_m = \cos^2(m\theta+\alpha)$ of reaching the state $|\widetilde{w}\rangle \in L$ after m iterations. For ℓ/N small, after $m = \frac{\pi}{4}\sqrt{N/\ell}$ iterations, the probability of reaching $|\widetilde{w}\rangle$ is close to 1.* □

The result (3.12) is consistent with Grover's original algorithm for single object search with $\ell = 1$, which has $m \approx \frac{\pi}{4}\sqrt{N}$.

Theorem 3.6. (Boyer, Brassard, Høyer and Tapp [3]). *Assume that ℓ/N is small. Then any search algorithm for ℓ objects, in the form of*

$$U_p U_{p-1} \dots U_1 |w_I\rangle,$$

where each $U_j, j = 1, 2, \dots, p$, is a unitary operator and $|w_I\rangle$ is an arbitrary combination state, takes in average $p = \mathcal{O}(\sqrt{N/\ell})$ iterations in order to reach the subspace L with a positive probability P independent of N and ℓ. Therefore, the generalized Grover algorithm in Cor. 3.5 is of optimal order.

Proof. This is the major theorem in [3]; see Section 7 and particularly Theorem 8 therein. Note also the work by C. Zalka who considered some measurement effects in [15]. □

Unfortunately, if the number ℓ of good items is not known in advance, Corollary 3.5 does not tell us when to stop the iteration. This problem was addressed in [3], and in another way in [4]. In a related context an equation arose that was not fully solved in [3]. We consider it in the final segment of this section. As in [3, §3], consider stopping the Grover process after j iterations, and, if a good object is not obtained, starting it over again from the beginning. From Corollary 3.5, the probability of success after j iterations is $\cos^2(j\theta - \alpha)$. By a well-known theorem of probability theory, if the probability of success in one "trial" is p, then the expected number of trials before success is achieved will be $1/p$. (The probability that success is achieved on the kth trial is $p(1-p)^{k-1}$. Therefore, the expected number of trials is

$$\sum_{k=1}^{\infty} kp(1-p)^{k-1} = -p \sum_{k=1}^{\infty} \frac{d}{dp}(1-p)^k = -p \frac{d}{dp} \frac{1-p}{p} \ , \tag{3.13}$$

which is $1/p$.) In our case, each trial consists of j Grover iterations, so the expected number of iterations before success is

$$E(j) = j \cdot \sec^2(j\theta - \alpha) \ .$$

The optimal number of iterations j is obtained by setting the derivative $E'(j)$ equal to zero:

$$\begin{aligned} 0 = E'(j) &= \sec^2(j\theta - \alpha) + 2j\theta \sec^2(j\theta - \alpha) \tan(j\theta - \alpha), \\ 2j\theta &= -\cot((j\theta - \alpha)) \ . \end{aligned} \tag{3.14}$$

(In [3, §3], this equation is derived in the form $4\vartheta j = \tan((2j+1)\vartheta)$, which is seen to be equivalent to (3.14) by noting that $\vartheta = \frac{\theta}{2} = \frac{\pi}{2} - \alpha$. Those authors then note that they have not solved the equation $4\vartheta j = \tan((2j+1)\vartheta)$ but proceed to use an ad hoc equation $z = \tan(z/2)$

with $z = 4\vartheta j$ instead.) Let us now approximate the solution j of (3.14) iteratively as follows. From (3.14),

$$\begin{aligned} &2j\theta \sin(j\theta - \alpha) + \cos(j\theta - \alpha) = 0 \ , \\ &e^{2i(\theta j - \alpha)} = (i2\theta j + 1)/(i2\theta j - 1) \ , \end{aligned} \tag{3.15}$$

and by taking the logarithm of both sides, we obtain

$$2i(\theta j - \alpha) = 2i\pi n + i \arg\left(\frac{i2\theta j + 1}{i2\theta j - 1}\right) + \ln\left|\frac{i2\theta j + 1}{i2\theta j - 1}\right| \ , \tag{3.16}$$

for any integer n. Assume that ℓ/N is small so that j is large, but we are looking for the smallest such positive j. Note that the logarithmic term in (3.16) vanishes, and

$$\begin{aligned} \arg\left(\frac{i2\theta j + 1}{i2\theta j - 1}\right) &= -2\tan^{-1}\frac{1}{2\theta j} \\ &= 2\left[\sum_{q=0}^{\infty} \frac{(-1)^{q+1}}{2q+1}\left(\frac{1}{2\theta j}\right)^{2q+1}\right] \\ &= -\frac{1}{\theta j} + \mathcal{O}((\theta j)^{-3}) \ ; \end{aligned}$$

by taking $n = 0$ in (3.16), we obtain

$$\begin{aligned} j &= \frac{1}{2i\theta}\left[2i\alpha - i \cdot \frac{1}{\theta j} + \mathcal{O}((\theta j)^{-3})\right] \\ &= \frac{1}{\theta}\left[\alpha - \frac{1}{2\theta j} + \mathcal{O}((\theta j)^{-3})\right] \ . \end{aligned} \tag{3.17}$$

The first order approximation j_1 for j is obtained by solving

$$\begin{aligned} j_1 &= \frac{1}{\theta}\left(\alpha - \frac{1}{2\theta j_1}\right) \ , \\ j_1^2 - \frac{1}{\theta}\alpha j_1 + \frac{1}{2\theta^2} &= 0 \ , \\ j_1 &= \frac{1}{2\theta}(\alpha + \sqrt{\alpha^2 - 2}) \ . \end{aligned} \tag{3.18}$$

Higher order approximations j_{n+1} for $n = 1, 2, \ldots$, may be obtained by successive iterations

$$j_{n+1} = \frac{1}{\theta}\left(\alpha - \tan^{-1}\frac{1}{2\theta j_n}\right)$$

based on (3.14). This process will yield a convergent solution j to (3.14).

Appendix: Random Multi-Object Search

Given a set of N unsorted objects, among which ℓ of them are the desired objects that we are searching for, how many times in average do we need to search in order to obtain the first desired object?

Because the N objects are unsorted, the above problem is equivalent to a familiar problem in probability theory:

> "An urn contains N balls, ℓ of them are black and the rest are white. Each time draw a ball randomly without replacement. Determine the number of times in average needed in order to draw the first black ball". (A.1)

A different version of random multi-object search would correspond to (A.1) but *with replacement* after each drawing. This search method is less efficient than (A.1), but can be treated by a similar approach as given below and, thus, will be omitted.

Even though we believe the solution to (A.1) is available in the literature, we could not locate a precise citation and, therefore, feel the urge to write this Appendix.

We define a random variable

$$\begin{aligned} T_b \equiv\ & \text{number of drawings needed to draw a ball randomly} \\ & \text{and without replacement until the first black ball is out.} \end{aligned} \tag{A.2}$$

Our objective is to calculate $E(T_b)$, the expected or the average value of T_b.

Obviously,

$$T_b \in \{1, 2, \ldots, N - \ell + 1\}. \tag{A.3}$$

We use $\binom{n}{j}$ to denote the combinatorial coefficient $n!/[j!(n-j)!]$, and use $P(A)$ to denote the probability of a given event A (measurable with respect to the random variable T_b).

Proposition A.1. *For $j \in \{1, 2, \ldots, N - \ell + 1\}$,*

$$P(T_b = j) = \frac{\binom{N-\ell}{j-1}}{\binom{N}{j-1}} \cdot \frac{\ell}{N - (j+1)}. \tag{A.4}$$

Proof. By the very definition of T_b in (A.2), we know that

$$\begin{aligned} P(T_b = j) = P(&\text{the first } j-1 \text{ drawings result in } j-1 \text{ white balls, but} \\ &\text{the } j\text{-th drawing results in a black ball}). \end{aligned}$$

Therefore (A.4) follows. □

Proposition A.2. *([11, p. 54, (10), (11)]) For any $m, n \in \{0, 1, 2, \ldots\}$ and $m \le n$,*

$$\sum_{k=0}^{n-m} \binom{m+k}{m} = \sum_{j=m}^{n} \binom{j}{m} = \binom{n+1}{m+1}$$

Proof. By Pascal's formula,

$$\binom{j+1}{m+1} = \binom{j}{m+1} + \binom{j}{m}, \tag{A.5}$$

we have

$$\begin{aligned}
\sum_{j=m}^{n} \binom{j}{m} &= \sum_{j=m}^{n} \left[\binom{j+1}{m+1} - \binom{j}{m+1} \right] \\
&= \sum_{j=m}^{n} \binom{j+1}{m+1} - \sum_{j=m-1}^{n-1} \binom{j+1}{m+1} \\
&= \binom{n+1}{m+1} - \binom{m}{m+1} = \binom{n+1}{m+1}. \quad \square
\end{aligned}$$

Theorem A.3.

$$E(T_b) = \frac{N+1}{\ell+1}. \tag{A.6}$$

Proof. From (A.4), we have

$$\begin{aligned}
E(T_b) &= \sum_{j=1}^{N-\ell+1} j \cdot \frac{\binom{N-\ell}{j-1} \cdot \ell}{\binom{N}{j-1} \cdot [N-(j-1)]} = \sum_{j=1}^{N-\ell+1} j \cdot \ell \cdot \frac{\frac{(N-\ell)!}{(j-1)!(N-\ell-j+1)!}}{\frac{N!}{(j-1)!(N-j+1)!} \cdot (N-j+1)} \\
&= \frac{\ell(N-\ell)!}{N!} \sum_{j=0}^{N-\ell} (j+1) \frac{(N-j-1)!}{(N-\ell-j)!} \\
&= \frac{\ell(N-\ell)!}{N!} \sum_{k=0}^{N-\ell} (N-\ell+1-k) \cdot \frac{(k+\ell-1)!}{k!} \quad (\text{where } k = N-\ell-j) \\
&= (N-\ell+1) \left[\frac{\ell(N-\ell)!}{N!} \sum_{k=0}^{N-\ell} \frac{(k+\ell-1)!}{k!} \right] - \frac{\ell(N-\ell)!}{N!} \sum_{k=1}^{N-\ell} \frac{(k+\ell-1)!}{(k-1)!} \\
&= (N-\ell+1) \left[\sum_{k=0}^{N-\ell} \frac{\frac{(k+\ell-1)!}{k!(\ell-1)!}}{\frac{N!}{\ell!(N-\ell)!}} \right] - \ell \left[\sum_{k=0}^{N-(\ell+1)} \frac{\frac{(k+\ell)!}{k!\ell!}}{\frac{N!}{\ell!(N-\ell)!}} \right] \\
&= (N-\ell+1) \left[\frac{1}{\binom{N}{\ell}} \sum_{k=0}^{N-\ell} \binom{k+\ell-1}{\ell-1} \right] - \ell \left[\frac{1}{\binom{N}{\ell}} \sum_{k=0}^{N-(\ell+1)} \binom{k+\ell}{\ell} \right].
\end{aligned} \tag{A.7}$$

Now, applying Proposition A.2, we obtain

$$\binom{N}{\ell-1} + \sum_{k=0}^{N-\ell} \binom{k+\ell-1}{\ell-1} = \binom{N}{\ell}, \tag{A.8}$$

$$\binom{N}{\ell} + \sum_{k=0}^{N-(\ell+1)} \binom{k+\ell}{\ell} = \binom{N}{\ell+1} = \frac{N-\ell}{\ell+1} \binom{N}{\ell}. \tag{A.9}$$

Substituting (A.8) and (A.9) into (A.7), we obtain

$$\begin{aligned}
E(T_b) &= \cdots (\text{continuing from (A.7)}) \\
&= (N-\ell+1) \binom{N}{\ell}^{-1} \binom{N}{\ell-1} - \ell + (N-\ell+1) - \ell \cdot \frac{N-\ell}{\ell+1} = \frac{N+1}{\ell+1}. \quad \square
\end{aligned}$$

Remark A.1. When $\ell = 1$, by (A.6) it takes $(N+1)/2$ searches in average to obtain the desired single object. In the literature, this is usually cited as $N/2$, which of course differs negligibly for large N.

For $N = 4$ and $\ell = 1$, by (A.6) it takes $(4+1)/(1+1) = 2.5$ times of search on average to obtain the desired item. But in [5, p. 3408], it is stated that it takes $9/4 = 2.25$ times of search on average. The reason for the discrepancy is a different definition of successful search. The authors of [5] regard the search as completed as soon as the location of the desired item is known, even if that item has not been physically "drawn from the urn". They therefore count the worst case, where the desired item is the last one drawn, as requiring only 3 steps instead of 4. This redefinition could be incorporated into our theorem at the expense of some complication; but it seems to us to be the less natural convention in the scenario of multiple desired objects only one of which is required to be "produced".

□

Acknowledgments: We thank M.O. Scully for originally expounding [6], B.-G. Englert and M. Hillery for acquainting us with some of the literature of quantum computation, M. M. Kash for a technical discussion, and Hwang Lee, D. A. Lidar, and J. D. Malley for comments on the manuscript.

References

[1] E. Biham, O. Biham, D. Biron, M. Grassl, and D.A. Lidar, Grover's quantum search algorithm for an arbitrary initial amplitude distribution, Phys. Rev. A **60** (1999), 2742–2745.

[2] D. Biron, O. Biham, E. Biham, M. Grassl, and D.A. Lidar, Generalized Grover search algorithm for arbitrary initial amplitude distribution, in *Quantum Computing and Quantum Communications* Lecture Notes. Comp. Sci. vol. **1509**, Springer, New York, 1998, pp. 140–147.

[3] M. Boyer, G. Brassard, P. Høyer and A. Tapp, Tight bounds on quantum searching, Fortsch. Phys. **46** (1998), 493–506.

[4] G. Brassard, P. Høyer and A. Tapp, Quantum counting, `quant-ph/9805082`, May 1998.

[5] I.L. Chuang, N. Gershenfeld and M. Kubinec, Experimental implementation of fast quantum searching, Phys. Rev. Lett. 80 (1998), 3408–3441.

[6] E. Farhi and S. Gutmann, Analog analogue of a digital quantum computation, Phys. Rev. **A 57** (1998), 2403–2405.

[7] L.K. Grover, A fast quantum mechanical algorithm for database search, Proc. 28th Annual Symposium on the Theory of Computing, 212–218, ACM Press, New York, 1996.

[8] L.K. Grover, Quantum mechanics helps in searching for a needle in a haystack, Phys. Rev. Letters **78** (1997), 325–328.

[9] L.K. Grover, Quantum computers can search rapidly by using almost any transformation, Phys. Rev. Letters **80** (1998), 4329–4332.

[10] R. Jozsa, Searching in Grover's algorithm, `quant-ph/9901021`, Jan. 1999.

[11] D. Knuth, *The Art of Computer Programming, Vol. 1, Fundamental Algorithms*, second edition, Reading, MA, 1973.

[12] A. Messiah, *Quantum Mechanics*, Vol. 2, Wiley, New York, 1966.

[13] P. Shor, Algorithms for quantum computation: discrete logarithms and factoring, Proc. 35th IEEE Symposium on the Foundations of Computer Sci., 124–134, 1994.

[14] P. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM J. Comp. **26** (1997), 1484–1510.

[15] C. Zalka, Grover's quantum searching algorithm is optimal, `quant-ph/9711070`, Nov. 1997.